

SLOVENSKÁ ŠTATISTIKA a DEMOGRAFIA

SLOVAK STATISTICS
and DEMOGRAPHY

2/2015

ročník/volume 25

Recenzovaný vedecký časopis so zameraním na prezentáciu moderných štatistických a demografických metód a postupov.

Scientific peer-reviewed journal focusing on the presentation of modern statistical and demographic methods and procedures.

Článok/Article: 4

Typ článku/Type of article: vedecký článok/scientific article

Strany/Pages: 43 – 57

Dátum vydania/Publication date: 15. apríl 2015/April 15, 2015



Pavol ĎURČEK

Prírodovedecká fakulta Univerzity Komenského v Bratislave

BILANCIA POHYBU OBYVATEĽSTVA PODĽA OBCÍ V ROKU 1971 – ÚPRAVA NA ÚROVEŇ EXISTUJÚCICH OBCÍ V ROKU 2013

THE BALANCE OF THE POPULATION CHANGE AT THE MUNICIPAL LEVEL IN 1971 – ADJUSTMENT TO THE EXISTING MUNICIPAL LEVEL IN 2013

ABSTRAKT

Tento príspevok má niekoľko parciálnych cieľov. Prvým a najzákladnejším je zdigitalizovanie hodnôt bilancie pohybu obyvateľstva v obciach Slovenska z roku 1971. Ďalším cieľom je dosiahnuť prepojitelnosť a porovnateľnosť dát z roku 1971 so súčasným územnosprávnym členením, predovšetkým s úrovňou obcí. Konečným cieľom je poukázať na možnú stratu informácie pri úprave dát z roku 1971 do súčasného obdobia. Na dosiahnutie posledného cieľa použijeme vybrané geoštatistické metódy.

ABSTRACT

This paper has several partial objectives. The first and the most important objective is the digitization of values of the population movement in the municipalities of Slovakia in 1971. Another objective is to achieve connectivity and comparability of data from 1971 with the present administrative division, especially with the level of municipalities. The final objective is to highlight the possible data loss due to data transfer from 1971 to the present. For the achievement of the last objective selected geostatistical methods will be used.

KLÚČOVÉ SLOVÁ

bilancia pohybu obyvateľstva, digitalizácia, porovnateľnosť dát, úroveň obcí

KEY WORDS

balance of the population change, digitalization, comparability of data, the municipal level

1. ÚVOD

Jedným z každoročne vypracúvaných demografických zdrojov je bilancia pohybu obyvateľstva za úroveň obcí. Vydávanie tohto dátového súboru na území SR sa začalo v roku 1971. Od tohto roku sú za úroveň obcí dostupné dáta týkajúce sa počtu obyvateľov na konci a začiatku roka plus základné údaje demografickej dynamiky, ako je počet narodených, počet zomretých, počet prisťahovaných a počet vystťahovaných. Bilancie pohybu obyvateľova za úroveň obcí sa však v období rokov 1971 až 1990 vydávali len v tlačenej podobe, čo pre súčasné obdobie nie je postačujúce. Na to, aby bola možná analýza a následná porovnateľnosť dát v dlhších časových intervaloch, je nevyhnutná ich digitalizácia. Na našom modelovom príklade sa pokúsime tieto problémy vyriešiť.

Naše ciele vychádzajúce z týchto potrieb sú: zdigitalizovať bilanciu pohybu obyvateľstva za úroveň obcí v roku 1971, realizovať prepojitelnosť a porovnateľnosť s údajmi z roku 2013, následne poukázať na možnú chybu, ktorá pri realizovaní prepojitelnosti mohla vzniknúť.

2. METODIKA

Prvým cieľom je digitalizácia. Digitalizácia analógových materiálov prebehla technologicky síce jednoduchším, no o to prácnejším spôsobom. Jednotlivé údaje sa z tlačných hárkov ručne prepisovali do tabuľkového súboru formátu xls. Pre jednotlivé obce sme prepisovali dáta v nasledujúcej štruktúre.

Tabuľka č. 1: Ukážka štruktúry prepisovaných dát

Kód obce	Názov obce	Pohlavie	Stav k 1.1.	Narod.	Zom.	Prist'. Vyst'.	Vyst'.	Stav k 31. 12.
505679	Kostolná Ves	muži	240	4	6	0	18	220
		ženy	210	0	1	3	11	201
		spolu	450	4	7	3	29	421

Zdroj: vlastné spracovanie

Žltou farbou sú vyznačené dáta, ktoré sa prepisovali. Konkrétne bolo prepísané číslo obce a názov obce platné v roku 1971. Ďalej sa prepisovali dáta za obe pohlavia samostatne. Konkrétne počet obyvateľov k 1. 1., počet narodených, počet zomretých, počet prisťahovaných a počet vystťahovaných. Zelenou farbou sú vyznačené dáta, ktoré sa následne na základe prepísaných dát dopočítavali. Hodnoty „spolu“ predstavovali sumu prepisovaných dát za mužov a ženy. Hodnoty pre „stav k 31. 12.“ predstavujú bilanciu. Tu sa k počtu obyvateľov k 1.1. pripočítali hodnoty počtu narodených, prisťahovaných a odpočítali hodnoty počtu zomretých a vystťahovaných. Zároveň dáta vyznačené zelenou farbou slúžili na kontrolu správnosti prepisovaných dát. Pri dopočítavaných hodnotách sa skontrolovala zhoda s dátami uvedenými v analógovom zdroji. Kontrola prebiehala dvojakým spôsobom. Samostatne sa kontrolovala správnosť hodnôt pre riadok „spolu“ a samostatne sa kontrolovala správnosť hodnôt pre stĺpec „stav k 31. 12.“.

Druhý cieľ – realizovať porovnateľnosť s údajmi z roku 2013 – bol časovo menej náročný, ale postupovo o to zložitejší. Zásadným problémom pri dosahovaní tohto cieľa bol fakt, že počet a vymedzenie obcí v roku 1971 nie je do značnej miery v zhode s počtom a vymedzením obcí v roku 2013. V roku 1971 bol počet obcí 3 085 (nerátame mestské časti Bratislavy a Košíc). V roku 2013 bol počet obcí 2 890 (nerátame mestské časti Bratislavy a Košíc). V roku 2013 malo Slovensko 116 obcí, ktoré v roku 1971 neexistovali. Naopak, v roku 2013 neexistovalo 251 obcí z roku 1971. Z týchto čísel vyplývajú dva problémy, a to problém agregácie dát a problém dezagregácie dát. Problém agregácie sa dá riešiť pomerne jednoducho. Dáta za zaniknuté obce môžeme pripočítať obciam, ku ktorým sa zaniknuté obce pričlenili alebo s ktorými sa zlúčili. Väčší problém predstavuje 116 obcí, ktoré v roku 1971 neexistovali.

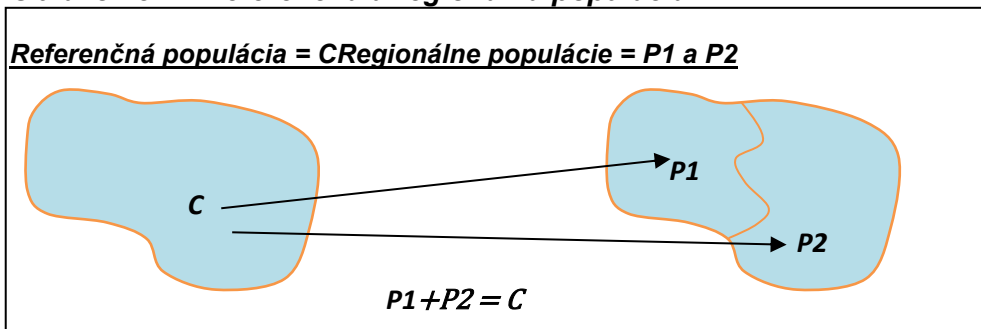
V roku 1971 bolo 116 neexistujúcich obcí súčasťou vtedajších 56 obcí. Delenie 56 obcí prebehlo do roku 2013 týmto spôsobom: každá jedna z 53 obcí z roku 1971 bola do roku 2013 rozdelená na 2 obce, každú z ďalších dvoch obcí, ktoré existovali v roku 1971, tvorili v roku 2013 tri obce a jedna obec z roku 1971 pozostávala v roku 2013 zo štyroch obcí. Prehľad obcí, ktoré neboli v roku 1971 súčasťou platnej územnosprávnej štruktúry, poskytuje tabuľka č. 3. Obsahuje aj informáciu o tom, do ktorých obcí boli v roku 1971 začlenené.

Proces rozdeľovania/dezagregovania dát prebiehal takto: Údaje bilancie pohybu obyvateľstva za 56 obcí z roku 1971 sme mali k dispozícii. Údaje o bilancii sme mohli pre každú z týchto 56 obcí rozdeliť na príslušný počet častí (na 116 obcí existujúcich v roku 2013) v istom pomere. Pomer, ktorým sme každú z 56 obcí delili, predstavuje počet obyvateľov neexistujúcich 116 obcí. Tento počet obyvateľov za neexistujúcich 116 obcí sme získali z dát za úroveň obcí zo sčítaní 1970 a 1980. Následne sme údaje za rok 1971 vypočítali pomocou modelu aritmeticky rastúceho regionálneho podielu, inak nazývaného aj SHIFT [3]. Matematický zápis modelu je

$$\hat{P}_{n+h} = C_{n+h} \left[\frac{P_n}{C_n} + \frac{h}{n} \left(\frac{P_n}{C_n} - \frac{P_0}{C_0} \right) \right]$$

V našom prípade tvorí referenčnú populáciu (C) každá z 56 existujúcich obcí a regionálnu populáciu (P) predstavuje každá zo 116 neexistujúcich obcí. Označenie premenných P a C matematického zápisu lepšie ilustruje obrázok č. 1.

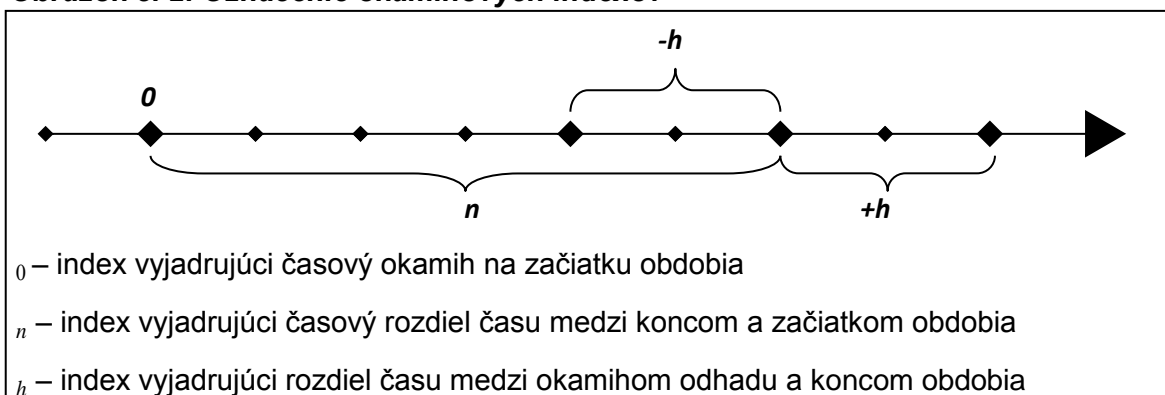
Obrázok č. 1: Referenčná a regionálna populácia



Zdroj: spracované na základe Rogerson, 2001

Dolné indexy matematického zápisu modelu SHIFT naznačujú rôzne časové okamihy, resp. časové úseky. Inak sa tieto indexy nazývajú aj časové indexy. Rovnako aj tu uvádzame jednoduchý nákres, ktorý na časovej osi vysvetľuje označenia jednotlivých časových indexov.

Obrázok č. 2: Označenie okamihových indexov



Zdroj: spracované na základe Rogerson, 2001

V našom prípade predstavuje začiatok obdobia dátum 1. 12. 1970. Koniec obdobia predstavuje dátum 1. 11. 1980. Dátum, ku ktorému chceme odhadovať, je

1. 7. 1971. Časový rozdiel medzi koncom a začiatkom obdobia je 9,92 roka (resp. 3623 dní). Časový rozdiel medzi okamihom odhadu a koncom obdobia je teda –9,35 roka (resp. –3 411 dní).

Ako názorný príklad celého procesu dezagregácie môžeme uviesť obce Turnianska Nová Ves, Chorváty a Host'ovce. V roku 2013 tieto obce existovali ako samostatné, avšak v roku 1971 tvorili obec Nová Bodva. Nová Bodva mala na základe dát bilancie pohybu obyvateľstva 970 obyvateľov k 1. 7. 1971. Na základe dát zo sčítania 1970 a 1980 a pomocou odhadovacej metódy SHARE sme zistili, že k 1. 7. 1971 majú obce Turnianska Nová Ves, Chorváty a Host'ovce počet obyvateľov 369, 209 a 391. Preto obec Turnianska Nová Ves predstavovala 38,09 % obyvateľov z pôvodnej obce Nová Bodva, obec Chorváty predstavovala 21,55 % obyvateľov z pôvodnej obce Nová Bodva a obec Host'ovce tvorila 40,36 % obyvateľov z pôvodnej obce Nová Bodva. Na základe tohto percentuálneho zastúpenia sme rozdelili aj ostatné dáta týkajúce sa bilancie pohybu obyvateľstva. Názorný príklad dezagregácie je schematicky načrtnutý v obrázku č. 3.

Obrázok č. 3: Schematický postup procesu dezagregácie dát

Bilancia pohybu obyvateľstva pre obec Nová Bodva existujúcu v roku 1971 (digitalizované dáta)						
Názov obce 1971	Stav k 1.1.	Narod.	Zom.	Prist'.	Vyst'.	Stav k 31.12.
Nová Bodva	964	11	10	21	10	976
Odhadnutý počet obyvateľov a následný pomer delenia ďalších dát pre obce neexistujúce v roku 1971						
Názov obce 1971	Stav 1.7.	Názov obce 2013	Stav 1.7.	Pomer delenia		
Nová Bodva	970	Turnianska Nová Ves	369,4	0,3808		
		Chorváty	209,1	0,2155		
		Host'ovce	391,5	0,4035		
Odhadnutá bilancia pohybu obyvateľstva pre obce neexistujúce v roku 1971						
Názov obce 2013	Stav k 1.1.	Narod.	Zom.	Prist'.	Vyst'.	Stav k 31.12.
Turnianska Nová Ves	367	4	4	8	4	372
Chorváty	208	2	2	5	2	210
Host'ovce	389	5	4	8	4	394

Zdroj: vlastné spracovanie

Na základe opísaných agregáčnych a dezagregačných postupov sme údaje o bilancii za úroveň obcí z roku 1971 prispôbili vymedzeniu obcí podľa platného územnosprávneho usporiadania SR v roku 2013. Tým sme dosiahli jeden z našich zámerov upraviť dáta tak, aby mohli byť porovnateľné so súčasným územnosprávnym členením.

Posledným cieľom nášho príspevku je zhodnotenie chyby, ktorá mohla vzniknúť prispôbením dát o bilancii za úroveň obcí z roku 1971 na úroveň obcí platnú v roku 2013. Na dosiahnutie tohto cieľa použijeme systém geoštatistických postupov. Prostredie, v ktorom sa realizovalo získavanie geopriestorových dát potrebných na určenie stupňa chybovosti pri prevode dát z úrovni obcí existujúcich v roku 1971 na úroveň obcí v roku 2013, predstavuje GIS balík ArcGis 10.2. Samotný test stupňa

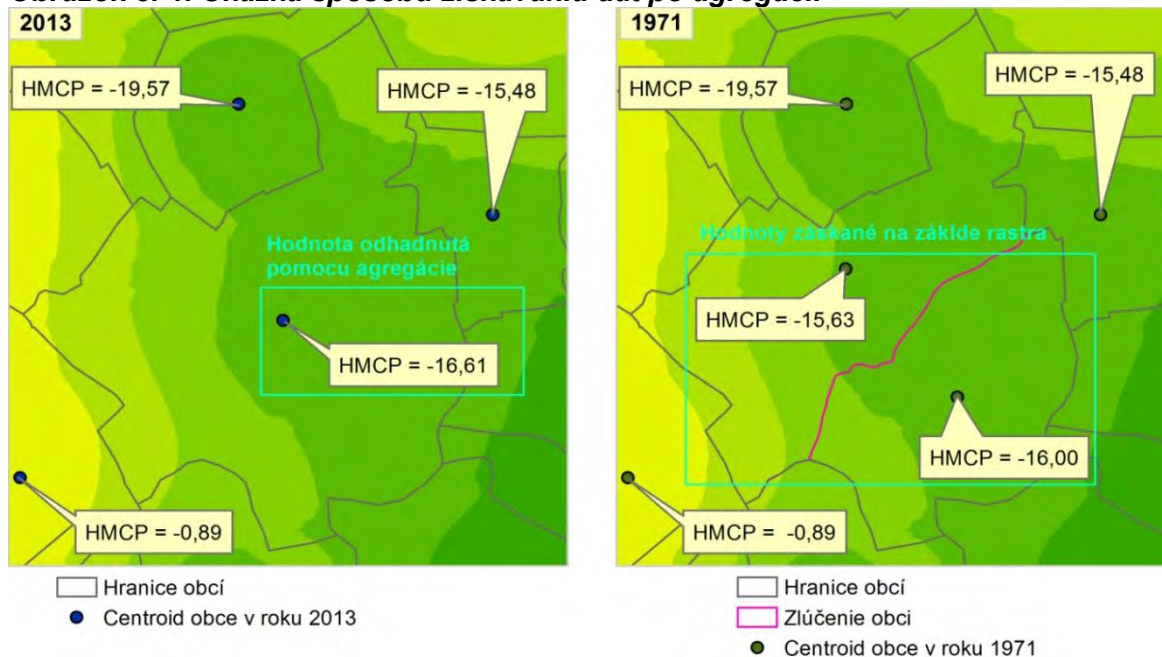
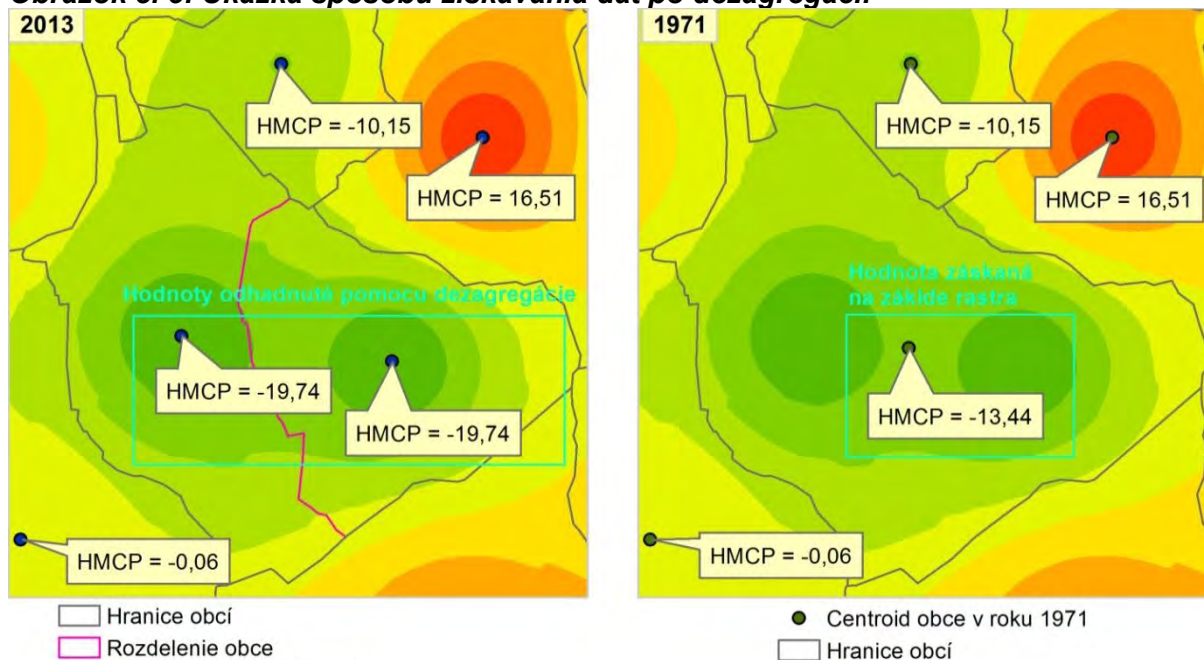
chybovosti bol realizovaný pomocou koeficientu determinácie, ktorý je druhou mocninou Pearsonovho korelačného koeficientu (označenie R^2). Koeficient determinácie vyjadruje, na koľko percent sa na seba podobajú oba analyzované štatistické súbory [2]. Na testovanie presnosti sme použili ukazovateľ hrubá miera celkového prírastku/úbytku (ďalej len „HMCP“ výsledky v promilových bodoch). Tento ukazovateľ v sebe zahŕňa všetky digitalizované údaje bilancie pohybu obyvateľstva.

Princíp hodnotenia stupňa chybovosti spočíva v tom, že porovnávame dva rastrové výstupy HMCP. Prvý výstup predstavuje raster, ktorý vznikol z dát a vymedzenia obcí v roku 1971. Druhý raster vznikol zobrazením dát po prevode údajov z roku 1971 na obecnú štruktúru platnú v roku 2013.

Pri tvorbe rastrov sa používa metóda IDW dostupná v rámci skupiny nástrojov Raster Interpolation. Rastre majú rozlíšenie 100 x 100 metrov. Vzniknuté rastre sme zobrazili a pomocou funkcie MINUS v rámci skupiny funkcií Raster Math sme oba rastre odčítali a tým sme získali priestorový obraz toho, kde sú rozdiely v HMCP najväčšie. To v podstate predstavuje akúsi priestorovú interpretáciu chyby, ktorá mohla pri prevode dát na obecnú štruktúru platnú v roku 2013 vzniknúť.

Ďalším krokom je celkové zhodnotenie chybovosti pomocou koeficientu determinácie. Tu sme postupovali tak, že z oboch rastrových výstupov (pre raster vytvorený z pôvodnej databázy a raster vytvorený z upravenej databázy) sme cez nástroj Extract Values To Points extrahovali hodnoty ukazovateľa HMCP pre centroidy (centroid predstavuje geometrický stred nejakého územia) obecnej štruktúry existujúcej v roku 1971. Tieto hodnoty sme následne podrobili hodnoteniu rozdielnosti. Princíp, ktorým sa dáta z rastrov získavali, naznačujú dva nasledujúce obrázky.

Extrahovanie hodnôt do centroidov obcí existujúcich v roku 1971 prinieslo tri rôzne výsledky. Prvú, najpočetnejšiu skupinu predstavujú centroidy obcí, ktoré sa nachádzali v roku 1971 aj v roku 2013 na rovnakom mieste. Tieto obce medzi rokmi 1971 až 2013 nezaznamenali zmenu obecných hraníc, a preto hodnoty HMCP budú pre oba varianty totožné. Na nasledujúcich obrázkoch predstavujú tieto obce centroidy, ktoré majú na pravom aj ľavom zobrazení hodnotu HMCP rovnakú. Ďalší prípad predstavujú obce, ktoré v roku 1971 existovali a do roku 2013 zanikli. Tento prípad reprezentuje obrázok 4. Hodnoty vzniknuté po agregácii boli extrahované do centroidov obecnej štruktúry platnej v roku 1971. Obrázok 5 následne reprezentuje posledný, tretí prípad, keď medzi rokmi 1971 až 2013 došlo k rozdeleniu obce. Hodnoty vzniknuté po dezagregácii boli extrahované do centroidov obecnej štruktúry platnej v roku 1971. Pre dva posledné varianty platí, že hodnoty extrahované z oboch rastrov nebudú totožné.

Obrázok č. 4: Ukážka spôsobu získavania dát po agregácii**Zdroj : vlastné spracovanie****Obrázok č. 5: Ukážka spôsobu získavania dát po dezagregácii****Zdroj : vlastné spracovanie**

Dáta z rôznych rastrov pre rovnaký súbor centroidov sme exportovali do súboru formátu xls. V takomto súbore sme následne pomocou funkcií PERSON a POWER vypočítali Pearsonov korelačný koeficient a koeficient determinácie. Ďalej sme analyzované dáta zobrazili v grafe XY závislosti. Pomocou koeficientu determinácie a rastrových mapových zobrazení sme realizovali hodnotenie globálne, ako aj hodnotenie priestorovej chybovosti.

3. TVORBA DATABÁZY A ÚPRAVA NA ÚROVEŇ OBCÍ EXISTUJÚCICH V ROKU 2013

Na základe postupu, ktorý je opísaný v metodike, sme realizovali digitalizáciu údajov bilancie pohybu obyvateľstva za úroveň obcí za rok 1971. Zároveň sme hodnoty za úroveň obcí spájali aj do vyšších územných jednotiek, ako sú okresy a kraje vymedzené v roku 1971 a okresy a kraje vymedzené v roku 1996. Za tieto vyššie územné jednotky sme zachovali pôvodnú štruktúru dát s výnimkou vystaňovaných a prisťahovaných. Počty prisťahovaných a vystaňovaných nie je možné agregovať z nižších územných jednotiek na vyššie. Výsledok agregácie počtu migrantov (prisťahovaných alebo vystaňovaných) by nedával reálny obraz o úrovni migrácie v danej vyššej územnej jednotke, pretože výsledok takejto agregácie by v sebe zahŕňal aj migrantov vo vnútri vyššej územnej jednotky, ako aj migráciu medzi samotnými vyššími územnými jednotkami. Z tohto dôvodu sme z dostupných dát vypočítali iba migračné saldo (to v sebe vnútornú migráciu vo vyšších územných jednotkách nezahŕňa). V tabuľke č. 2 je ukážka štruktúry dát pre vyššie územné jednotky.

Tabuľka č. 2: Ukážka štruktúry dát pre vyššie územné jednotky

Kód okresu/kraja	Názov okresu/kraja	Pohlavie	Stav k 1. 1.	Narod.	Zom.	Saldo sťahov.	Stav k 31. 12.
------------------	--------------------	----------	--------------	--------	------	---------------	----------------

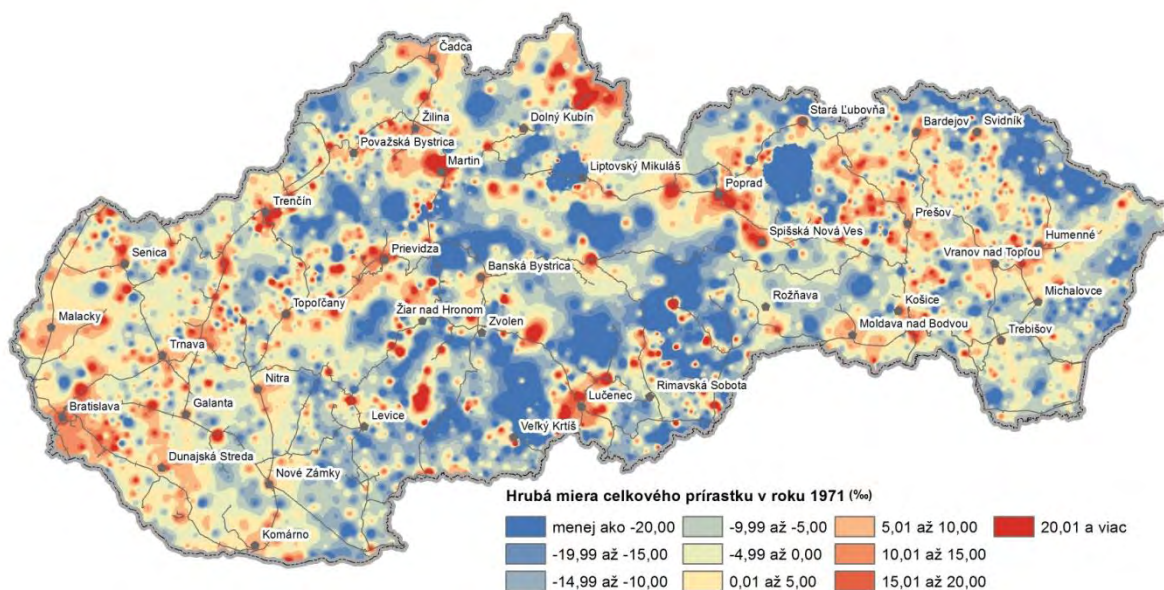
Zdroj: vlastné spracovanie

Následne sme digitalizované dáta za úroveň obcí v roku 1971 upravili na úroveň obcí existujúcich v roku 2013. Vzniknutú databázu sme uložili vo formáte xls. na internetovej adrese, kde je voľne dostupná na stiahnutie: <https://docs.google.com/file/d/0B6n4KynlTPZcQzJYUi11Z1VTRTA/edit>

4. HODNOTENIE ÚROVNE CHYBOVOSTI PO PREVODE NA ÚROVEŇ OBCÍ V ROKU 2013

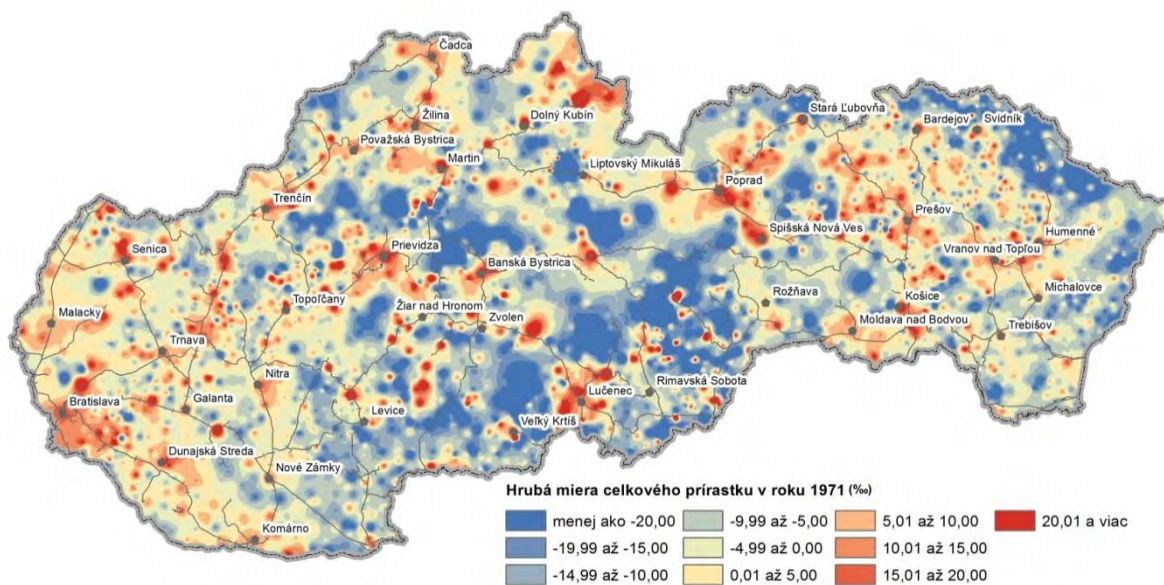
Z metodiky vyplýva, že sme porovnávali priestorový (rastrový) obraz vytvorený z reálnych dát za rok 1971 a priestorový obraz vytvorený z dát upravených na obecnú štruktúru platnú za rok 2013. Na nasledujúcich dvoch obrázkoch môžeme vidieť rastrové zobrazenie HMCP v roku 1971 pre obecnú štruktúru existujúcu v roku 1971 a obecnú štruktúru existujúcu v roku 2013. V obrázku č. 8 je jednoduchým odčítaním rastrov znázornený absolútny rozdiel medzi oboma rastrami. Pomocou týchto mapových zobrazení môžeme pomerne jednoducho identifikovať regióny, v ktorých došlo v dôsledku prevodu dát na obecnú štruktúru platnú v roku 2013 k najväčším informačným stratám.

Obrázok č. 6: HMCP v roku 1971 zobrazená na základe obecnej štruktúry platnej v roku 1971

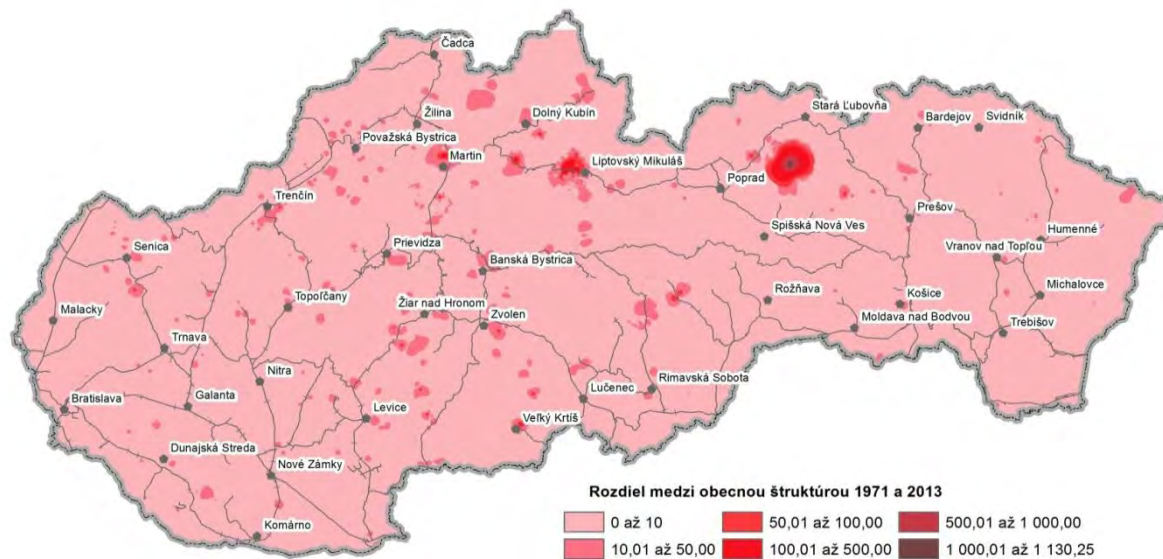


Zdroj: vlastné spracovanie

Obrázok č. 7: HMCP v roku 1971 zobrazená na základe obecnej štruktúry platnej v roku 2013



Zdroj: vlastné spracovanie

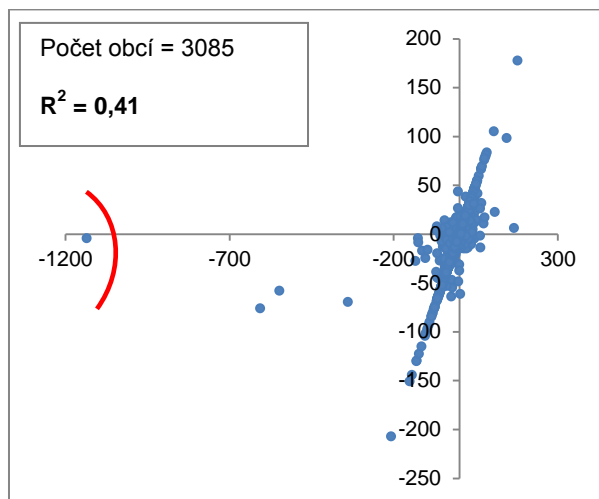
Obrázok č. 8: Absolútny rozdiel HMCP pre obecné štruktúry platné v rokoch 1971 a 2013

Zdroj: vlastné spracovanie

Na obrázku č. 8 vidno, že najväčšie rozdiely v hodnotách HMCP medzi rôznymi obecnými štruktúrami sa nachádzajú predovšetkým v oblasti Levočských vrchov, v okolí priehrad Nová Bystrica, Liptovská Mara, Starina a v priestore miest a tzv. strediskových obcí. Dôvody zaniknutých rozdielov spočívajú v tom, že v oblasti Levočských vrchov sa v roku 1971 nachádzal vojenský obvod Javorina, ktorý bol v roku 2012 zrušený a jeho územie bolo porozdeľované medzi susedné obce. To spôsobilo výraznú zmenu územia jednotlivých obcí a tým aj výraznú zmenu polohy centroidov týchto obcí. Ďalej výstavba veľkých vodných diel urýchlila zánik viacerých obcí. Územia týchto zaniknutých obcí boli následne pričlenené k okolitým obciam, čo sa prejavilo na zmene polohy centroidov. V prípade miest a strediskových obcí nastala medzi rokmi 1971 a 2013 výrazná zmena rozlohy ich územia v dôsledku pričleňovania alebo odčleňovania blízkych obcí. To sa tiež prejavovalo na výraznej zmene polohy centroidov miest a strediskových obcí.

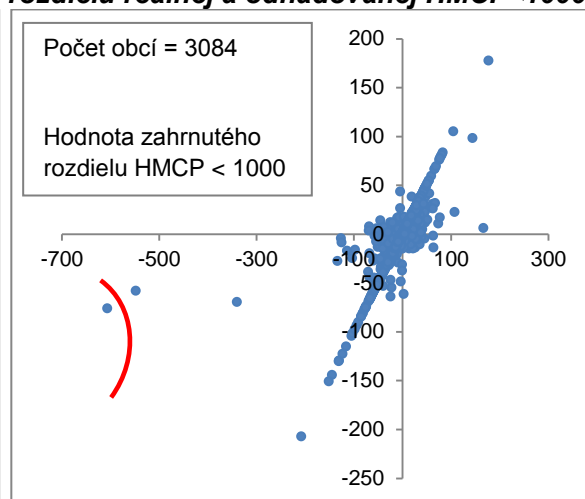
Ďalej sme pomocou dát extrahovaných do centroidov obcí hodnotili chybovosť, ktorá mohla pri prevode dát na inú obecnú štruktúru vzniknúť. Výsledky sme sa rozhodli znázorniť v nasledujúcej sérii korelačných grafov, v ktorých sme zobrazili aj koeficient determinácie (R^2). Nijaký graf okrem prvého nezahŕňa skupinu obcí, ktoré sa vyznačovali najvyššími rozdielmi HMCP (viac opis v samotnom grafe). V grafoch je zároveň červenou čiarou naznačené, ktorá skupina obcí nebude v nasledujúcom grafe zaradená do korelačnej analýzy.

Graf č. 1: Korelácia pre celý súbor obcí

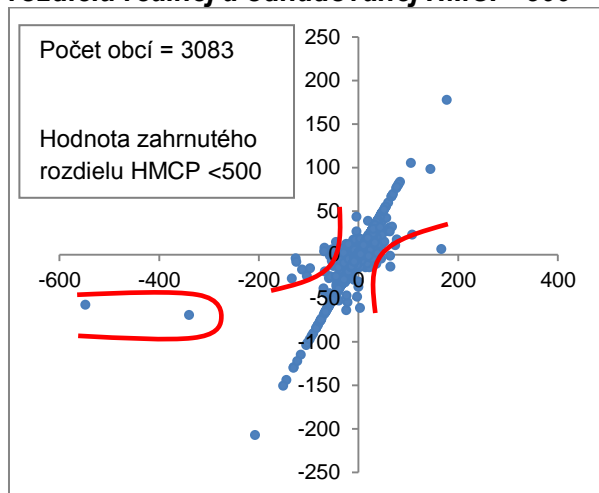


Zdroj: vlastné spracovanie

Graf č. 2: Korelácia pre obce s hodnotou rozdielu reálnej a odhadovanej HMCP < 1000

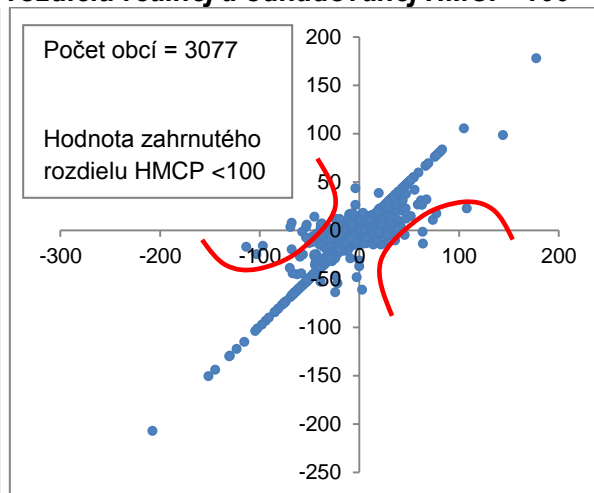


Graf č. 3: Korelácia pre obce s hodnotou rozdielu reálnej a odhadovanej HMCP < 500

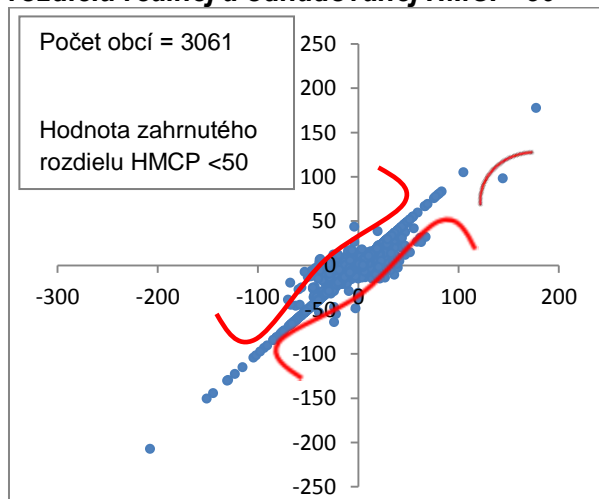


Zdroj: vlastné spracovanie

Graf č. 4: Korelácia pre obce s hodnotou rozdielu reálnej a odhadovanej HMCP < 100

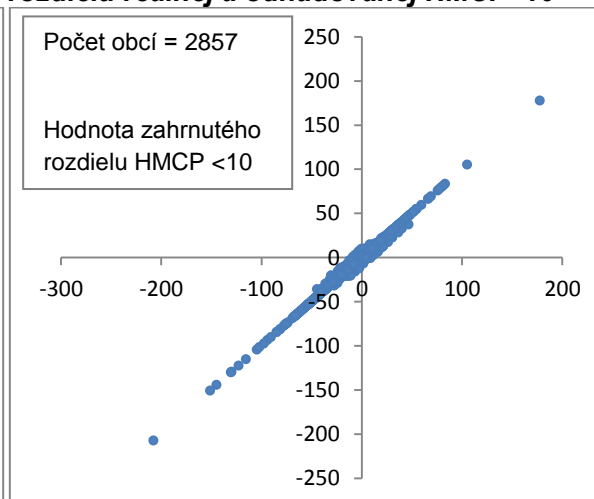


Graf č. 5: Korelácia pre obce s hodnotou rozdielu reálnej a odhadovanej HMCP < 50



Zdroj: vlastné spracovanie

Graf č. 6: Korelácia pre obce s hodnotou rozdielu reálnej a odhadovanej HMCP < 10



Z grafov vyplýva, že koeficient determinácie pre celý súbor predstavuje iba 0,41, čo sa môže interpretovať tak, že reálne dáta a odhadnuté dáta sa na seba podobajú na 41 %. Ak z analyzovaného štatistického súboru odstránime 8 obcí s najväčším rozdielom HMCP (hodnota rozdielu je viac ako 100), tak nám koeficient determinácie stúpne na 0,85, resp. na 85 %. Na základe obrázkov č. 6 až 8 môžeme tieto obce identifikovať v priestore zaniknutého vojenského obvodu Javorina. Následne po vyradení obcí s rozdielom HMCP väčším ako 10 (ide o 128 obcí) dostávame hodnotu koeficientu determinácie väčšiu ako 0,99, resp. 99 %. Tieto obce sa nachádzajú hlavne v lokalitách, kde došlo k výraznej zmene obecných hraníc, čiže hlavne v priestoroch vznikajúcich vodných diel alebo v okolí miest a strediskových obcí, kde dochádzalo k spájaniu alebo rozdeľovaniu obcí.

Namiesto je otázka, či má takto redukovaný súbor dostatočnú výpovednú hodnotu a či použitý odhadový postup možno označiť za relevantný a dostatočne presný. Graf 6 naznačuje, že 99-percentnú presnosť odhadu sme dosiahli vtedy, keď sme do korelačnej analýzy nezahrnuli 128 obcí s najväčším rozdielom HMCP. Týchto 128 obcí predstavuje 4,15 % z celkového počtu obcí (populačná váha týchto obcí z celkovej populácie je ešte menšia, a to 0,09 %). Preto môžeme povedať, že pre viac ako 95 % obcí je odhad a prevod dát relevantný.

5. ZÁVER

Zámerom tohto článku bolo poukázať na význam digitalizácie dát týkajúcich sa bilancie pohybu. Globálnym cieľom bolo zdigitalizovať dáta o bilancii pohybu obyvateľstva obcí v roku 1971 a zabezpečiť porovnateľnosť týchto dát s územnosprávnym členením v roku 2013. Proces digitalizácie bol časovo náročný, ale z hľadiska technickej náročnosti pomerne jednoduchý, na rozdiel od procesu prevedenia dát o bilancii z obecnej štruktúry existujúcej v roku 1971 na obecnú štruktúru v roku 2013, ktorý bol časovo menej náročný, ale metodicky pomerne zložitý. Výsledkom je databáza dostupná v pôvodnej podobe, ako aj v podobe platnej na úroveň obcí existujúcich v súčasnosti.

Naším parciálnym cieľom bolo hodnotenie chybovosti dát po ich prevode na obecnú štruktúru platnú v roku 2013. Celý súbor vykazoval pomerne značnú chybovosť, pretože koeficient determinácie dosiahol hodnotu 0,44. Ak sme však z hodnotenia chybovosti vyradili 5,6 % obcí s najväčšími rozdielmi reálnych a odhadovaných dát, tak sme dosiahli takmer 100 % zhodu dátových súborov (koeficient determinácie mal hodnotu 0,99).

Pridanú hodnotu tohto príspevku predstavujú hlavne metodické postupy a predovšetkým vytvorená databáza. Táto databáza je zverejnená na príslušnom internetovom odkaze a môže voľne slúžiť všetkým, ktorí majú záujem dané dáta ďalej šíriť, analyzovať a hodnotiť.

Príloha**Obce neexistujúce v roku 1971 – porovnanie administratívnej štruktúry v rokoch 1971 a 2013**

Kód obce 1971	Názov obce 1971	Kód obce 2013	Názov obce 2013
508438	Banská Bystrica	508438 + 580244 + 557277 + 557285	Banská Bystrica + Malachov + Kynceľová + Nemce
514519	Bátka	514519 + 557919	Bátka + Dulovo
502065	Beša	502065 + 556777	Beša + Jesenské
501042	Bodza	501042 + 555819	Bodza + Bodzianske Lúky
505871	Bošáca	505871 + 556424	Bošáca + Haluzice
511340	České Brezovo	511340 + 582051	České Brezovo + Zlatno
501557	Dolný Bar	501557 + 555649	Dolný Bar + Mad
521345	Družstevná pri Hornáde	521345 + 582514	Družstevná pri Hornáde + Kostofany nad Hornádom
501565	Dvorníky na Ostrove	555665 + 555673	Malé Dvorníky + Veľké Dvorníky
542890	Haláčovce	542890 + 556289	Haláčovce + Otrhánky
501590	Horná Potôň	501590 + 582522	Horná Potôň + Potônske Lúky
501603	Horný Bar	501603 + 503461	Horný Bar + Bodíky
521434	Hraničná pri Hornáde	559687 + 580252	Kechnec + Milhošť
521541	Hutníky	559831 + 559865	Bočiar + Sokolany
501654	Jahodná	501654 + 555541	Jahodná + Dunajský Klátov
511498	Kokava nad Rimavicou	511498 + 580317	Kokava nad Rimavicou + Utekáč
520411	Krásny Brod	520411 + 559610	Krásny Brod + Rokytovce
518549	Kriváň	518549 + 580520	Kriváň + Korytárky
501719	Kútники	501719 + 555720	Kútники + Povoda
500445	Lapáš	555851 + 555860	Malý Lapáš + Veľký Lapáš
511218	Lučenec	511218 + 580309	Lučenec + Mikušovce
508811	Mýto pod Ďumbierom	508811 + 557251	Mýto pod Ďumbierom + Bystrá
505170	Nadlice	505170 + 556173	Nadlice + Livina
524913	Nemcovce	524913 + 559971	Nemcovce + Lipníky
500615	Nitrany	555886 + 555908	Čechynce + Malý Cetín
521779	Nová Bodva	518107 + 599310 + 521779	Turnianska Nová Ves + Chorváty + Hostovce
520560	Ohradzany	520560 + 559636	Ohradzany + Sopkovce
501859	Orechová Potôň	501859 + 555746	Orechová Potôň + Vieska
500674	Podhorany	500674 + 582697	Podhorany + Bádice
543462	Poľanovce	543462 + 581640	Poľanovce + Korytné
524140	Prešov	524140 + 518522 + 518590	Prešov + Haniska + Ľubotice
501336	Radvaň nad Dunajom	501336 + 555827	Radvaň nad Dunajom + Virt
515388	Ratkovská Lehota	515388 + 557820	Ratkovská Lehota + Sása
503983	Reca	503983 + 582549	Reca + Nový Svet
504751	Rohovské Rybky	556122 + 556131	Rohov + Rybky
525529	Rožňava	525529 + 560031	Rožňava + Čučma
528714	Ruská	528714 + 513857	Ruská + Budince
520772	Slovenská Volová	520772 + 559644	Slovenská Volová + Gruzovce

dokončenie			
Kód obce 1971	Názov obce 1971	Kód obce 2013	Názov obce 2013
509019	Staré Hory	509019 + 557269	Staré Hory + Turecká
543802	Streda nad Bodrogom	543802 + 513831	Streda nad Bodrogom + Klin nad Bodrogom
503584	Štúrovo	503584 + 556092	Štúrovo + Nána
501948	Trhové Mýto	555568 + 555576	Horné Mýto + Trhová Hradská
543837	Tŕňa	513792 + 513806	Malá Tŕňa + Veľká Tŕňa
527939	Turany nad Ondavou	527939 + 560073	Turany nad Ondavou + Vyšný Hrabovec
522112	Turnianske Podhradie	518123 + 559784	Háj + Turňa nad Bodvou
515701	Uzovská Panica	515701 + 557854	Uzovská Panica + Rakytník
515736	Veľké Teriakovce	515736 + 557790	Veľké Teriakovce + Nižný Skálnik
505731	Veľký Klíž	505731 + 556416	Veľký Klíž + Ješkova Ves
501999	Vojka nad Dunajom	501999 + 580597	Vojka nad Dunajom + Kyselica
544051	Vranov nad Topľou	544051 + 581674	Vranov nad Topľou + Hencovce
500941	Výčapy-Opatovce	500941 + 581097	Výčapy-Opatovce + Ľudovítová
521035	Vyšný Hrušov	521035 + 559652	Vyšný Hrušov + Maškovce
522236	Zádielske Dvorníky	559873 + 559881	Dvorníky-Včeláre + Zádiel
544001	Zemplínska Nová Ves	544001 + 513849	Zemplínska Nová Ves + Stanča
502022	Zlaté Klasy	502022 + 580554	Zlaté Klasy + Čenkovce
524131	Ždiar	524131 + 580368	Ždiar + Tatranská Javorina

Zdroj: Štatistický úrad SR [4], [5], [6], [7], Jazykovedný ústav Ľ. Štúra SAV [1]

Príspevok je súčasťou riešenia grantového projektu VEGA č. 1/0562/12 Nové demografické analýzy a prognózy obyvateľstva Slovenska a jeho regiónov s využitím progresívnych geografických aplikácií.

POĎAKOVANIE

Podakovanie autora článku patrí študentom 3. ročníka bakalárskeho programu geografia a demografia na Prírodovedeckej fakulte Univerzity Komenského v Bratislave v akademickom roku 2012/2013, ktorí sa významne podieľali na digitalizácii dát o bilancii pohybu obyvateľstva obcí v roku 1971. Bez odbornej a svedomitej práce každého študenta by nebolo možné tento príspevok zostaviť.

LITERATÚRA

- [1] MAJTÁN, M.: Názvy obcí Slovenskej republiky. Vývin v rokoch 1773 – 1997. Bratislava: Veda, 1998. Dostupné na: <http://slovníky.juls.savba.sk/>
- [2] NOVÁKOVÁ, G.: Štatistika pre geografov. Bratislava: Geografika, 2013. 227 s. ISBN 978-80-89317-18-9.
- [3] ROGERSON, P.: Statistical methods for geography. London: SAGE publications, 2001. 232 s. ISBN 0-7619-6287-5.
- [4] ŠÚ SR: Bilancia obyvateľstva obcí v roku 1971. 1972 (analogový zdroj).
- [5] ŠÚ SR: Bilancia pohybu obyvateľstva podľa obcí v roku 2013. 2014.
- [6] ŠÚ SR: Štatistický lexikón obcí Slovenskej republiky 1980. Bratislava: ŠÚ SR, 1982.

[7] ŠÚ SR: Štatistický lexikón obcí Slovenskej republiky 1970. Bratislava: ŠÚ SR, 1973.

RESUMÉ

Cieľom príspevku bolo upozorniť na digitalizáciu dát o bilancii pohybu obyvateľstva za obce z roku 1971 a proces jej tvorby. Digitalizovanú databázu sme ďalej upravili na úroveň obcí v roku 2013. Následne sme zhodnotili chybu, ktorá mohla pri tejto úprave dát nastať.

Na dosiahnutie týchto cieľov sme zostavili pomerne zložitý algoritmus jednotlivých pracovných etáp. Po digitalizácii časti dát sme následne ostatné údaje dopočítavali. Na základe dopočítavaných dát sa realizovala kontrola správnosti dát.

Po zdigitalizovaní a skontrolovaní dát sme realizovali náš ďalší cieľ, ktorým bola úprava dát na obecnú štruktúru existujúcu v roku 2013. V roku 1971 bol počet obcí 3 085 (nerátame mestské časti Bratislavy a Košíc). V roku 2013 sa ich počet znížil na 2 890 obcí (nerátame mestské časti Bratislavy a Košíc). Oproti roku 2013 neexistovalo v roku 1971 116 obcí. Naopak, v roku 2013 neexistovalo 251 obcí z roku 1971. Z uvedeného vyplýva, že 87 % obcí nezaznamenalo žiadnu zmenu svojich hraníc. Zvyšných 13 % obcí prešlo agregáciou alebo dezagregáciou, na základe ktorých boli hodnoty odhadnuté.

Po prevode dát na obecnú štruktúru roku 2013 sme následne realizovali hodnotenie chybovosti tohto prevodu. Hodnotenie sa uskutočnilo dvojakým spôsobom. Prvý spôsob predstavuje priestorové vymedzenie areálov, v ktorých bola rozdielnosť reálnych a odhadovaných dát najväčšia. Tu sme použili jednoduchý princíp, keď sme od seba odčítali dve rastrové zobrazenia. Prvý raster predstavoval hodnoty HMCP v roku 1971 pre obecnú štruktúru existujúcu v roku 1971 a druhý hodnoty HMCP v roku 1971 po procese úpravy na obecnú štruktúru existujúcu v roku 2013. Druhý spôsob spočíval v hodnotení oboch rastrov pomocou koeficientu determinácie.

Výsledkom našej práce je databáza, ktorá obsahuje dáta o bilancii pohybu obyvateľstva za obce v roku 1971. Táto databáza je dostupná za pôvodnú obecnú štruktúru roku 1971, ako aj obecnú štruktúru roku 2013. Ďalej sa tieto dáta jednoduchým spôsobom sčítali do vyšších územných celkov, okresov a krajov.

Databáza je dostupná na:

<https://docs.google.com/file/d/0B6n4KynITPZcQzJYUi11Z1VTRTA/edit>

Globálna chybovosť porovnávanej skutočnej a upravenej databázy bola pomerne výrazná, pretože zhoda vypočítaná koeficientom determinácie predstavovala iba 44 %. Po odstránení 5,6 % obcí s najväčším rozdielom reálnych a odhadovaných dát sme dosiahli prakticky zhodu dátových súborov (koeficient determinácie mal hodnotu 99 %).

RESUME

The main objective of this paper was to stress the implementation of data digitization of the population movement in the municipalities from year 1971. The digitized database was adjusted to the level of municipalities in year 2013. We subsequently evaluated the possible errors caused by adjustment of data.

To achieve these goals a complex algorithm of each work stage was established. After data digitization the rest of the data was grossed up. On the basis of these data the data correctness was performed.

Once digitization and data checking were finished, our next objective was realized, which was the adjustment of data to the structure of municipalities in 2013. In 1971, the number of municipalities was 3085 (if we exclude the city parts of Bratislava and

Košice). In 2013, the number of municipalities was reduced to 2890 (if we exclude the city parts of Bratislava and Košice). In 1971 compared with 2013, the number of municipalities was 116. On the contrary,, in 2013 there weren't 251 like in 1971. As mentioned above, 87 % municipalities did not observed any border changes. The remaining, 13 % were subjected to the process of aggregation and disaggregation, what on the basis of which estimation of values was carried out.

After the data transfer to the structure of municipalities in 2013, the evaluation of the error rate of this transfer was realized. This evaluation was carried out in two ways. The first way represents the spatial definition of the premises in which the nonconformity was the highest. A simple principle was used here when two raster displays were deduced from each other. The first raster represented the HMCP values in 1971 for the structure of municipalities and the second raster the HMCP values in 1971 after making adjustments to the structure of municipalities in 2013. The second evaluation method was based on the assessment of both rasters provided by the coefficient of determination.

As a result, a database was created, containing the data on balance of population movement according to municipalities in 1971. This database is available in original structure of municipalities in 1971 and for year 2013. These data were further summarized into higher territorial units like districts and regions.

The database is available through the following link:

<https://docs.google.com/file/d/0B6n4KynlTPZcQzJYUi11Z1VTRTA/edit>

The global error rate of the compared actual and modified database was quite significant, since the conformity calculated by the coefficient of determination was only 44 %. After the removal of 5.6 % of the municipalities with the largest local discrepancy nearly identical data files were obtained (coefficient of determination was equal to 99 %).

PROFESIJNÝ ŽIVOTOPIS

Mgr. Pavol Ďurček vyštudoval odbor geografia a demogeografia v štátnej správe a samospráve na Prírodovedeckej fakulte Univerzity Komenského v Bratislave. Od roku 2011 pôsobí na Katedre humánnej geografie a demografie ako doktorand. Venuje sa výskumu geografickej, ako aj štatistickej diferenciacie demografických procesov a demografických štruktúr.

KONTAKT

durcek@fns.uniba.sk